IFW AF

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appellants:	Krishna A. BHARAT et al.	§	Confirmation No.:		8878
		§			
Serial No.:	09/418,418	§	Group Art Unit:		2172
		§			
Filed:	10/15/1999	§	Examiner:		B. N. To
		§			
For:	Method For Ranking	§	Docket No.:		200308296-1
	Hypertext Search Results	§			
	By Analysis Of Hyperlinks	§			
	From Expert Documents	§			
	And Keyword Scope	§			

APPEAL BRIEF

Mail Stop Appeal Brief – Patents   Date: October 13, 2004
Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

Appellants hereby submit this Appeal Brief in connection with the above-identified application. A Notice of Appeal was filed via facsimile on August 13, 2004.

Appl. No. 09/418,418
Appeal Brief dated October 13, 2004
Reply to final Office action of June 29, 2004

TABLE OF CONTENTS

I.   REAL PARTY IN INTEREST

The real party in interest is the Hewlett-Packard Development Company
(HPDC), a Texas Limited Partnership, having its principal place of business in
Houston, Texas, through its merger with Compaq Computer Corporation (CCC)
which owned Compaq Information Technologies Group, L.P. (CITG).   The
assignment from the CCC to CITG was recorded on November 16, 2001, at
Reel/Frame 012305/0944. The Change of Name Document was recorded on
May 12, 2004, at Reel/Frame 014628/0103.

II.   RELATED APPEALS AND INTERFERENCES

Appellants are unaware of any related appeals or interferences.

III. **STATUS OF THE CLAIMS**

Originally filed claims: 1-21.

Claim cancellations: 13

Added claims: 22-23.

Presently pending claims: 1-12 and 14-23.

Presently appealed claims: 1-12 and 14-23.

IV.   STATUS OF THE AMENDMENTS

No claims were amended after the final Office action dated June 29, 2004.

V.   SUMMARY OF THE CLAIMED SUBJECT MATTER

The invention is a solution to a problem caused by the fact that the WWW contains an extremely large number of documents. Conventional search engines based on searching for a given topic may require excessive amounts of time and computer resources to actually search the entire content of the web to identify documents related to a submitted topic when that topic is submitted. They also have difficulty in filtering out documents which are of low quality. See page 1, lines 9-15.

The present invention reduces the time and effort required to provide a ranked list of documents after a query is submitted by breaking the process into two broad phases. See page 4, lines 1-2. The first phase, expert lookup, is performed before any search query is received, i.e., in a preprocessing step. See page 4, lines 1-8, page 6, lines 23-24, and step 202 in Fig. 2(a). The second phase is a topic-based search which occurs only after receipt of a topic based query and requires searching of only the expert documents from the first phase and documents to which the experts point. See page 4, lines 9-17, page 6, lines 25-29 and steps 206, 208 of Fig. 2(a). The search after receipt of a topic-based query is dramatically simplified by reducing the number of documents which must be searched.

Figs. 3(a) and 3(b) provide alternative methods for determining expert pages which is the first phase of the invention. These methods are described in the specification beginning at page 8, line 25 through page 9, line 9. The list of expert pages is created without reference to a topic. The process for identifying experts is based on factors other than specific topics.

The first phase may also include indexing of the expert list by topics to create an expert reverse index. See Fig. 4 and the specification at page 9, line 11 through page 10, line 9. This step also occurs after the experts have been identified, but before a topic based query is received.

The second phase begins when a topic-based query is received. It begins by ranking the expert documents according to the submitted topic. See Fig. 8 and the specification at page 10, line 11, through page 11, line 19. As with any

known topic based ranking, documents which do not have any, or not enough, key phrases matching the query topic receive a low or zero ranking, which is equivalent to being deleted or removed from the subset created by the ranking process. This step therefore results in a small set of expert documents related to the query topic.

The second phase then includes ranking target documents identified by the small set of expert documents identified above. See Fig. 9 and the specification at page 11, line 21 through page 13, line 2. The target document set is dramatically smaller than the set of all documents on the web which may include the query topic. But the target documents are qualified by having been identified, pointed to, by the expert documents identified in the first phase of the process. As with known topic based ranking processes, the documents which receive low scores are deleted and not included in the list returned in response to the query.

VI.   GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Whether claims 1, 20 and 21 are patentable under 35 U.S.C. 112, second paragraph.

Whether claims 1, 20 and 21-23 are patentable under 35 U.S.C. 103(a) over Chakrabarti et al. (Automatic resource compilation by analyzing hyperlink structure and associated Text April 14, 1998) in view of Liddy et al. (U.S. Patent No. 6,304,864).

Whether claims 2-10 and 19 are patentable under 35 U.S.C. 103(a) over Chakrabarti et al. in view of Liddy et al. and Yu (U.S. Patent No. 6,167,552).

Whether claims 14-18 are patentable under 35 U.S.C. 103(a) over Chakrabarti et al. in view of Liddy et al. and Page (U.S. Patent No. 6,285,999).

Whether claims 11-12 are patentable under 35 U.S.C. 103(a) over Chakrabarti et al in view of Liddy et al and Chakrabarti (U.S. Patent No. 4,418,433).

VII.   ARGUMENT

A.   Detailed argument re 112 rejections.

In the final Office action of June 29, 2004, the independent claims 1, 20 and 21 were for the first time rejected under 35 U.S.C. 112, second paragraph, as being incomplete for omitting essential steps, such omission amounting to a gap between the steps. The allegedly omitted steps are: searching step for the topic after the crawling process and before the ranking of the expert documents in accordance with the search query.

The Appellants submit that claims 1, 20 and 21 clearly point out and distinctly claim the subject matter which the Appellants regard as the invention.

The first step of the second phase of the invention is ranking the expert documents (identified in the first phase) after receipt of a query topic. As noted in the summary of the invention above, it is well known in the art that topic based ranking means searching the content of a set of documents (in this case the set of expert documents) for the topic and applying various metrics to assign a relevancy value. Documents which do not contain the topic terms or otherwise fail to meet the metrics are assigned a low, or zero, relevancy value. Low relevancy documents are not reported back and may be considered to be deleted from the list of documents. Thus, it is inherent in the ranking step that the documents are searched for the topic, since that is in fact a primary metric by which a relevancy score is assigned.

B.   Detailed argument re 103 rejections.

Claims 1, 20 and 21-23 were finally rejected as being unpatentable over the Chakrabarti et al publication, (Automatic resource compilation by analyzing hyperlink structure and associated Text April 14, 1998) in view of the Liddy et al. (US Patent No. 6,304,864).

None of the references teaches or suggests the step of "forming a set of expert documents from the set of all hypertext documents crawled without reference to the search query." All of the references perform searching based on a query topic.

In the final Office action with specific reference to claims 1, 20 and 21, the Examiner acknowledged that "Chakrabarti does not teach forming a set of expert documents from the set of all hypertext documents crawled without reference to the search query." The Examiner asserts that "Liddy teaches 'the agents crawl the web and retrieve relevant documents (col. 4, lines 29-59) and rank the retrieval result by the values (col. 11, lines 28-39)." The Examiner further asserts that it would therefore have been obvious to include crawling and ranking the results as disclosed by Liddy into Chakrabarti in order to organize relevancy of documents on the web.

The Appellants agree that Chakrabarti does not teach forming a set of expert documents without reference to the query. The first step disclosed by Chakrabarti is a topic-based search of a collection of pages. In Section 2, Algorithm, Chakrabarti says that "Given a topic, the algorithm first gathers a collection of pages from which it will distill ones that it considers the best for the topic." Chakrabarti explains that the way to form this set is: "The topic is sent to a term-based search engine _ AltaVista in our case – and a root set of 200 documents containing the topic term(s) is collected." Thus Chakrabarti begins its process with a topic-based search. Chakrabarti could not begin the process without the topic, because the first step uses the topic.

The Liddy reference deals primarily with training of neural networks to crawl the web. The neural networks are trained to identify and rank relevant documents. However, relevancy is based on the query. The process begins with a user "query representing the information the user wishes to retrieve." See the Abstract. Then the agents execute searches on different search engines on the network "in accordance with the query." The "subject categories and important terms of each document" are processed to determine the retrieval value, i.e., relevancy, of each document. At Col. 6, lines 27-29, Liddy states that the agent server operates in accordance with the user profile, which includes the query, see col. 5, lines 66-67. The retrieval value of a document is based on the frequency or absence of the subject categories and important terms in the query. See col. 6, lines 50-53. Throughout the Liddy reference, every process for training a neural

network to search for documents, and every searching step begins with the terms from the query. Thus Liddy begins its process with a topic-based search. Liddy could not begin the process without the topic, because the first step uses the topic.
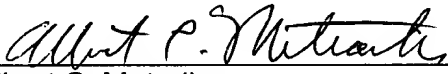
Thus neither of the cited references teaches or suggests forming a set of expert documents from the set of all documents crawled without reference to the search query. Since neither reference includes such a teaching, the combination cannot contain or make obvious such a teaching. The Appellants submit that Claims 1, 20 and 21 are patentable over the cited references.

All other appealed claims depend directly or indirectly from Claims 1, 20 and 21. Since Claims 1, 20 and 21 have been shown to be patentable over the cited references, the Appellants submit that the other appealed claims are also patentable over the cited references.

VIII.   CONCLUSION

For the reasons stated above, Appellants respectfully submit that the Examiner erred in rejecting all pending claims. It is believed that no extensions of time or fees are required, beyond those that may otherwise be provided for in documents accompanying this paper.   However, in the event that additional extensions of time are necessary to allow consideration of this paper, such extensions are hereby petitioned under 37 C.F.R. § 1.136(a), and any fees required (including fees for net addition of claims) are hereby authorized to be charged to Hewlett-Packard Development Company's Deposit Account No. 08-2025.

Respectfully submitted,

Albert C. Metrailer
PTO Reg. No. 27,145
CONLEY ROSE, P.C.
(713) 238-8000 (Phone)
(713) 238-8008 (Fax)
ATTORNEY FOR APPELLANTS

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
Legal Dept., M/S 35
P.O. Box 272400
Fort Collins, CO 80527-2400

IX.   CLAIMS APPENDIX

1.   (Previously presented) A computer-implemented method for searching a large number of hypertext documents in accordance with a search query, comprising:

>   forming a set of expert documents from the set of all hypertext documents
>   crawled without reference to the search query;
>   ranking the expert documents in accordance with the search query;
>   ranking target documents pointed to by the ranked expert documents; and
>   returning a results list based on the ranked target documents.

2.   (Original) The computer-implemented method of claim 1, wherein the hypertext documents are pages in the world wide web.

3.   (Original) The computer-implemented method of claim 1, wherein the hypertext documents are sites in the world wide web.

4.   (Original) The computer-implemented method of claim 1, wherein the hypertext documents are documents in a hypertext database.

5.   (Original) The computer-implemented method of claim 1, wherein an expert reverse index is constructed in memory for keywords appearing in the expert documents, the expert reverse index identifying the location of the keywords in the expert documents.

6.   (Original) The computer-implemented method of claim 5, wherein a keyword of an expert document is included in the expert reverse index if the keyword is part of a key phrase that qualifies at least one URL in the expert document.

7.   (Original) The computer-implemented method of claim 6, wherein a key phrase qualifies a URL if the URL is within the scope of the key phrase in the expert document.

8.   (Original) The computer-implemented method of claim 6, wherein a key phrase in an HTML title qualifies all URLs in the entire document.

9.   (Original) The computer-implemented method of claim 6, wherein a key phrase in an HTML heading qualifies all URLs in that portion of the document before a next HTML heading in the document of greater or equal importance.

10.   (Original) The computer-implemented method of claim 6, wherein a key phrase in an HTML anchor qualifies the URLs in the anchor.

11.   (Previously presented) The computer-implemented method of claim 1, wherein forming a set of expert documents includes:

    determining a document having at least a predetermined number of outlinks to be an expert document if the document also points to at least the predetermined number of targets on distinct non-affiliated hosts.

12.   (Original) The computer-implemented method of claim 11, wherein expert documents additionally must point to documents that share the same broad classification. `

13.   (Canceled).

14.   (Original) The computer-implemented method of claim 1, wherein ranking target documents pointed to by the expert documents includes:

determining a plurality of edge scores for each target document, where an edge score is determined for edges between the expert documents and the target document;

determining a target score in accordance with the edge scores of the target document;

ranking the target documents in accordance with the target scores.

15. (Previously presented) The computer-implemented method of claim 14, further including:

determining an edge score only for those links to the target document from a predetermined number of top-ranked expert documents.

16. (Original) The computer-implemented method of claim 14, further including selecting target documents to be ranked that are linked to by at least two mutually non-affiliated selected expert documents, where the selected target also is not affiliated with the expert documents.

17. (Original) The computer-implemented method of claim 14, where an edge score between an expert document and a target document ES(E,T) is determined as follows, where ExpertScore reflects the rankings of the expert documents:

a) find #occurrences of each keyword in all keyphrases of expert document E

b) if the #occurrences for any keyword in E is 0: ES(E,T)=0

else ES(E,T)=ExpertScore(E) * sum of #occurrences for all keywords.

18. (Original) The computer-implemented method of claim 14, wherein, if two affiliated experts have edges to the same target, the edge having a lower edge score is discarded and is not used to determine the target score.

19.   (Previously presented) The computer-implemented method of claim 18, wherein two hypertext documents are affiliated if at least one of the following is true: 1) they share the same rightmost non-generic suffix and 2) they have an IP address in common.

20.   (Previously presented) An apparatus that searches a large number of hypertext documents in accordance with a search query, comprising:

a software portion configured to form a set of expert documents from the set of all documents crawled without reference to the search query;

a software portion configured to rank the expert documents in accordance with the search query;

a software portion configured to rank target documents pointed to by the ranked expert documents; and

a software portion configured to return a results list based on the ranked target documents.

21.   (Previously presented) A computer program product, comprising:

a computer usable medium having computer readable instructions stored therein to search a large number of hypertext documents in accordance with a search query, including:

computer readable program code devices for causing a computer to form a set of expert documents from the set of all documents crawled without reference to the search query;

computer readable program code devices for causing a computer to rank the expert documents in accordance with the search query;

computer readable program code devices for causing a computer to rank target documents pointed to by the ranked expert documents; and

computer readable program code devices for causing a computer to return a results list based on the ranked target documents.

22.   (Previously presented) The computer-implemented method of claim 1, wherein ranking the expert documents in accordance with the search query comprises:

      determining a level score for each of the expert documents;

      determining a fullness factor for each key phrase on each of the expert documents; and

      determining an expert score for each expert document in accordance with the level score of the expert document and the fullness factors for the key phrases of the expert document.

23.   (Previously presented) The computer-implemented method of claim 1, forming a set of expert documents occurs before a search query is received.

ORIGINAL
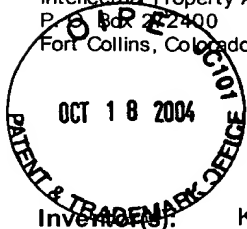
IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Krishna A. BHARAT et al.

Confirmation No.: 8878

Application No.: 09/418,418

Examiner: B. N. To

Filing Date: 10/15/1999

Group Art Unit: 2172

Title: METHOD FOR RANKING HYPERTEXT SEARCH RESULTS BY ANALYSIS OF HYPERLINKS FROM EXPERT DOCUMENTS AND KEYWORD SCOPE

Mail Stop Appeal Brief-Patents
Commissioner For Patents
PO Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL OF APPEAL BRIEF

Sir:

Transmitted herewith is the Appeal Brief in this application with respect to the Notice of Appeal filed on __08/13/2004__ .

The fee for filing this Appeal Brief is (37 CFR 1.17(c)) $340.00.

(complete (a) or (b) as applicable)

The proceedings herein are for a patent application and the provisions of 37 CFR 1.136(a) apply.

( ) (a) Applicant petitions for an extension of time under 37 CFR 1.136 (fees: 37 CFR 1.17(a)-(d) for the total number of months checked below:

( )	one month	$110.00
( )	two months	$430.00
( )	three months	$980.00
( )	four months	$1530.00

( ) The extension fee has already been filled in this application.

(X) (b) Applicant believes that no extension of time is required. However, this conditional petition is being made to provide for the possibility that applicant has inadvertently overlooked the need for a petition and fee for extension of time.

Please charge to Deposit Account **08-2025** the sum of __$340.00__ . At any time during the pendency of this application, please charge any fees required or credit any over payment to Deposit Account 08-2025 pursuant to 37 CFR 1.25. Additionally please charge any fees to Deposit Account 08-2025 under 37 CFR 1.16 through 1.21 inclusive, and any other sections in Title 37 of the Code of Federal Regulations that may regulate fees. A duplicate copy of this sheet is enclosed.

(X) I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, Alexandria, VA 22313-1450. Date of Deposit:__10/13/2004__
OR
( ) I hereby certify that this paper is being transmitted to the Patent and Trademark Office facsimile number_____ on _____

Number of pages:

Typed Name: Colleen F. Brown

Signature: _____

Rev 10/04 (Apbbrief)

Respectfully submitted,

Krishna A. BHARAT et al.

By _____

Albert C. Metrailer

Attorney/Agent for Applicant(s)
Reg. No. **27,145**

Date: **10/13/2004**

Telephone No.: **(713) 238-8000**
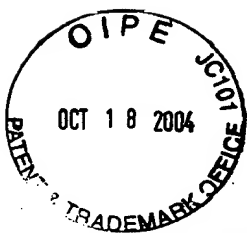
CHANGE OF CORRESPONDENCE ADDRESS *Application*	Application Number	09/418,418
	Filing Date	10/15/99
	First Named Inventor	Krishna Bharat
Address to:	Art Unit	2172
Commissioner for Patents P.O. Box 1450 Alexandria, VA 22313-1450.	Examiner Name	BaoQuoc N Tu
	Attorney Docket Number	200308296-1

Please change the Correspondence Address for the above-identified patent application to:

[X] Customer Number : | 022879 |

OR

[X] Firm or Individual Name Address	HEWLETT-PACKARD COMPANY Intellectual Property Administration (IPA)
Address	PO Box 272400
City	Fort Collins	State	CO	Zip	80527-2400
Country	USA
Telephone	970-898-3800	Fax	970-898-7247

This form cannot be used to change the data associated with a Customer Number. To change the data associated with an existing Customer Number use "Request for Customer Number Data Change" (PTO/SB/124).

I am the:

[ ] Applicant/Inventor

[ ] Assignee of record of the entire interest.
Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96).

[X] Attorney or Agent of record. Registration Number ___38,423___

[ ] Registered practitioner named in the application transmittal letter in an application without an executed oath or declaration. See 37 CFR 1.33(a)(1). Registration Number

Typed or Printed Name	Lloyd E. Dakin, Jr.
Signature	
Date	02-20-04	Telephone	650-857-2295

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.

[X] *Total of ___1___ forms are submitted.